DOCUMENT RESUME

ED 148 634 SE 023 786

AUTHOR Willson, Victor L.

TITLE Multifactor Analysis of Differences Between Correlation Coefficients. Research Paper No. 13. Revised Edition.

INSTITUTION Minnesota Univ., Minneapolis. Coll. of Education.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE Jan 75
GRANT NSF-GW-6800
NOTE 15p.; For related documents, see SE 023 784-792

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Academic Ability; *Achievement; *Educational Research; *Mathematics; *Research Methodology; *Science Education; Secondary Education; Secondary School Science; *Statistical Analysis; Student Attitudes
IDENTIFIERS Minnesota Research and Evaluation Project; National Science Foundation; *Research Reports

ABSTRACT

The derivation and analysis presented in this report allow the experimenter to perform multifactor contrasts on correlation coefficients for several populations in an experimental design. Cells of the design constitute independent groups of subjects on which two measures have been taken and correlated. Measures need not be identical for each group. An example is given in which the dependent variable represents the correlation between attitude and achievement for six secondary school science classes stratified by ability (three levels) and grade level (two levels). An overall relationship was observed between science attitude and knowledge of science for junior and senior high school students of science. No difference due to achievement was found, but the grade difference was significant. (Author/BB)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED148634

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

minnesota research & evaluation project college of education university of minnesota

funded by the national science foundation

SE 023 786

Research Paper #13

Multifactor Analysis of Differences Between Correlation Coefficients

Victor L. Willson
University of Minnesota
November 1974
Revised - January 1975

This study was supported by Grant GW6800 from the National Science Foundation to the University of Minnesota; Wayne W. Welch, Project Director.

Multifactor Analysis of Differences Between Correlation Coefficients

The general procedure for analysis of differences between correlation coefficients with blocking factors present is outlined. Cells of the design constitute independent groups of subjects on which two measures have been taken and correlated. Measures need not be identical for each group. An example is given in which the dependent variable represents the correlation between attitude and achievement for six secondary school groups, stratified by ability (three levels) and grade level (two levels).

A researcher may be interested in comparing correlation coefficients among several populations which have been sampled using a factorial design. Coefficients will normally be calculated for the same variables measured within each population. The researcher may be interested in the effects due to inclusion in specific levels of the design factors. Alternatively, the correlations calculated may be based upon different measures for each population, and the researcher may be interested in trends or differences between levels.

Previous Research

Since the theory of quadratic forms is necessary to the derivations in the paper, some background is needed. Box (1953) discussed the quadratic form

$$Q = z'Mz \qquad (1)$$

where $\underline{z}$ is a vector of normally distributed random variables with mean zero and variance-covariance diagonal matrix $\underline{\Sigma}$. $\underline{M}$ is a positive semi-definite matrix. Box showed Q is distributed as a linear sum of chi-square random variables

$$X = \sum_{i=1}^{I} \lambda_i \chi_1^2 \qquad (2)$$

where λ_i is the i-th latent root of the matrix

$$W = \Sigma M \tag{3}$$

for the characteristic equation

$$\det \left| \Sigma W - \lambda I \right| = 0. \tag{4}$$

The statistic

$$Q = \sum_{i=1}^{I} M_i \, (Z_i - Z.)^2 \tag{5}$$

is a quadratic form in which the latent roots are $\lambda_i = M_i \sigma_{z_i}^2$ for diagonal M and Σ matrices. Box showed that the distribution of Q is given approximately by

$$Q \sim g\chi^2(h) \quad \text{where} \tag{6}$$

$$g = \frac{\sum_{i=1}^{I} \lambda_i^2}{\sum_{i=1}^{I} \lambda_i} \quad \text{, and} \tag{7}$$

$$h = \frac{(\sum_{i=1}^{I} \lambda_i)^2}{\sum_{i=1}^{I} \lambda_i^2} \quad . \tag{8}$$

Among specific references to correlation comparisons, Hays (1963) gives the statistical procedure for testing the equivalence of J independent correlation coefficients. Marascuilo (1966) discussed the χ^2 analog of Scheffe's theorem for multiple comparisons among correlation coefficients r_k of K independent bivariate normal populations. The statistic he used is

$$U_o = \sum_{k=1}^{K} (n_k-3)\ (Z_{r_k} - Z_{r_o})^2 \quad , \qquad (9)$$

$$Z_{r_k} = 1/2\ \mathrm{Log}_e\ \frac{1+r_k}{1-r_k} \quad , \qquad (10)$$

$$Z_{r_o} = \frac{\sum_{k=1}^{K} (n_k-3)\ Z_{r_k}}{\sum_{k=1}^{K} (n_k-3)} \quad , \qquad (11)$$

where

n_k = sample size for k-th sample,

Z_{r_k} = Fisher - Z transformation of the correlations coefficient for the k-th sample, and

Z_{r_o} = common estimate of the k transformed correlations.

The U-statistic has an approximate χ^2 distribution with k-1 degrees of freedom. A derivation of this follows directly from Box's theorem on the approximate distribution of the quadratic form for the contrasts of (9).

Thomas and Stanley (1969) were interested in sex and race differences in correlations of variables predicting college freshman grade point average. They transformed the coefficients to Fisher Z-statistics and compared mean differences within sex and race. No formal model was developed, however. A formal statement of such a model seems in order.

A Statistical Model

Subjects are randomly assigned to levels of a multifactor experiment. Two measures are made on each subject and the measures correlated. The experimenter is interested in differences in the correlations due to the

factors and in possible interaction. Each cell in the design contains one "observation," the correlation coefficient. Since no within-cell variance exists for a single observation in such a layout, no independent error term can be specified ordinarily. Since variance errors are known, however, they will be utilized.

For a two-way layout, to be used as the basic example for the remainder of the paper, the Fisher-transformed correlation is decomposed into components:

$$Z_{r_{ij}} = Z_\rho + \alpha_i + \beta_j + \alpha\beta_{ij} + e_{ij} \quad , \qquad (12)$$

where the α, β and $\alpha\beta$ terms are factorial effects in the transformed correlations. Also, the standard error is $\sigma_{Z_{r_{ij}}} = 1 / \sqrt{n_{ij}-3}$ for the ij-th cell observation.

Under the usual constraints on the treatment parameters, the model represented by (12) is similar to the fixed effects ANOVA model. Normality holds, except in extreme situations (see Norris and Hjelm, 1960). Errors are assumed to be independent and within-cell variances are homogeneous if and only if $n_{ij} = n$ for all i and j.

One seeming advantage of working with a metric-free statistic such as the correlation coefficient is that coefficients can be compared directly, even though the original measures used to calculate the coefficients were different. This is the informal procedure used in construct validation. For example, correlations between attitude and achievement would necessarily be based on different instruments in comparing eight-year-old children and adults. Putting aside questions of unreliability of measurement, differences in correlation might provide results of theoretical interest, while direct comparisons of the original variables are difficult.

Equal Cell Sizes

When sample sizes used to compute the cell correlations are equal for all cells ($n_{ij}=n$), mean squares for effects may be calculated in the usual manner. It is then easy to show that (12) reduces under the null hypothesis to

$$U_{effect} = \frac{MS_{effect}}{(\frac{1}{n-3})} \sim \chi^2_{df_{effect}} . \qquad (13)$$

In the two-way layout

$$U_A = \frac{MS_A}{(\frac{1}{n-3})} \sim \chi^2_{I-1} , \qquad (14)$$

$$U_B = \frac{MS_B}{(\frac{1}{n-3})} \sim \chi^2_{J-1} , \qquad (15)$$

$$U_{AxB} = \frac{MS_{AxB}}{(\frac{1}{n-3})} \sim \chi^2_{(I-1)(J-1)} . \qquad (16)$$

These statistics may be shown to be Marascuilo's U-statistics, as well. The tests are independent, since the χ^2 statistics are independent. An additional test of interest may be made on the grand mean $Z_{r..}$,

$$z = \frac{Z_{r..} - Z_\rho}{(\frac{1}{IJ(n-3)})} , \qquad (17)$$

which is a normally distributed statistic. Thus, one can test the overall correlation for several populations.

Unequal Cell Sizes (Or Proportional Cell Sizes)

Analysis of the two-way crossed design is based on the known variance of each Fisher-transformed correlation. Box (1953) outlined the analysis for the nonorthogonal one-way design using the F-approximation for the ratio of two quadratic forms. Since the variances are not estimated but known for the transformed correlations (and linear combinations of them), the Box theorem given in (6) may be applied. Searle (1971) and Marks (1974) have discussed the major ways of approaching the nonorthogonal two-factor analysis of variance. The main effects model seems most amenable to solution and interpretation here.

The quadratic forms for the treatment affects in the two-way layout with unequal sample size in the cells under the null hypothesis

are $$Q_A = \sum_{i=1}^{I} m_i \left(Z_{r_{i.}} - Z_{r..}\right)^2 \quad , \qquad (21)$$

$$Q_B = \sum_{j=1}^{J} p_j \left(Z_{r_{.j}} - Z_{r..}\right)^2 \quad , \qquad (22)$$

where $$Z_{r_{i.}} = \sum_{j=1}^{J} (n_{ij}-3)\, Zr_{ij} \,/N_i \quad , \qquad (23)$$

$$Z_{r_{.j}} = \sum_{i=1}^{I} (n_{ij}-3)\, Zr_{ij} \,/N_j \quad , \qquad (24)$$

$$Zr.. = \sum\sum_{ij} (n_{ij}-3)\, Z_{r_{ij}} \,/N \quad , \qquad (25)$$

$$N_i = \sum_{j=1}^{J} (n_{ij}-3) \quad , \qquad (26)$$

$$N_j = \sum_{i=1}^{I} (n_{ij}-3) \quad , \text{ and} \qquad (27)$$

$$N = \sum\sum_{ij} (n_{ij}-3) \quad . \qquad (28)$$

Following Sheffe' (1959)

$$M_i = \left[\sum_{j=1}^{J} \frac{n_{ij}-3}{N_i^2} \right]^{-1} \qquad (29)$$

From Box (1954) The matrix W is

$$W = N^{-1} \left\{ \delta_{ij} \sigma_{i.} m_i N - \sigma_{i.}^2 m_i m_k \right\} \qquad (30)$$

when $\sigma_{i.}^2 = VAR \left[Z_{ri.} - Z_{r_{i..}} \right]$ (31)

$$= \frac{1}{\sum_{j=1}^{J} (n_{ij} - 3)}$$

The latent roots for W are then calculated from (30) and the approximation given in (9) is calculated, using one degree of freedom for each contrast.

A similar statistic for the effect Q_B may be calculated. Since h will generally not be integer-valued, the simplest procedure will be to interpolate between the nearest integers to find critical values.

The test of hypothesis is over the whole set of contrasts about the grand mean. Confidence intervals and contrasts based on the χ^2 analog to Scheffe's Theorem may be calculated, as noted by Marasciulo (1966).

The test for overall correlation is given by

$$z = \frac{Z_{r..} - Z_{\rho}}{\hat{\bar{\sigma}}_{ij}^2} \qquad (32)$$

where $\hat{\bar{\sigma}}_{ij}^2 = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{1}{n_{ij}-3} \Big/ IJ$

and z is a normally-distributed variable with mean zero and variance unity.

Since variances are unequal in the proportional design, the approach outlined here seems to be appropriate. Alternatively, one might use a procedure such as unweighted means analysis, if proportions are close to one.

Illustrative Example

As part of a National Science Foundation grant, Welch and Gullickson (1973) described a testing program in which randomly selected junior and senior high school classes in fifteen states were given a series of attitude and achievement measures in science and mathematics. Each science class was randomly divided into thirds, each third taking the Learning Attitude Inventory or LEI (Anderson, 1971), the Science Process Inventory or SPI (Welch and Pella, 1967), and the Test of Achievement in Science or TAS (Lawrenz, 1971). Two forms of the TAS were developed--Form I for the eighth grade level and Form II for the eleventh grade level. Items were drawn from the released items of the National Assessment of Educational Progress (1970). Kuder-Richardson reliabilities reported are .87 for both Forms I and II (Garibaldi, 1974). Since students generally did not take all tests, class means were used as the units of analysis.

Classes were divided into low, middle, and high thirds on the basis of TAS achievement scores, division performed separately for each Form. This produced a 2x3 factorial design with grade level and ability group as factors. Correlation coefficients were computed within-cell between the TAS and SAI with class means as unit of analysis. This correlation gives a measure of association between the class's attitude toward science and knowledge of science. Coefficient values and sample sizes are given in Table 1.

The analysis procedure for equal cell sizes was performed on the correlation coefficients. The U-statistics for effects are given in Table 2.

TABLE 1

Correlations Between SAI and TAS

for Two Factors, Achievement and Grade Level

Grade Level	Achievement Level			
	Low	Middle	High	Marginals
Junior High	r = .006	.247	.218	
(8th grade)	(n = 35)	(35)	(35)	
	Z_r = .006	.252	.222	$Z_{r_{1.}}$ = .160
Senior High	.401	.418	.390	
(11th grade)	(35)	(35)	(35)	
	Z_r = .425	.445	.412	$Z_{r_{2.}}$ = .427
Marginals	$Z_{r_{.1}}$ = .216	$Z_{r_{.2}}$ = .349	$Z_{r_{.3}}$ = .314	$Z_{r_{..}}$ = .294

Z_r = Fisher-transformed value

n = sample size in parentheses

TABLE 2

Statistics for Effects on Correlation

Between Achievement and Attitude

Effect	Statistic Distribution	d.f.	S.S.	Statistic Value	Probability Under Null Hypothesis
Grand Mean	Normal	1	-	56.45	< .01
Achievement	X^2	2	.019	.304	> .80
Grade	X^2	1	.107	3.45	< .07
Achievement x Grade	X^2	2	.0178	.570	> .7
Total	_____	6	.1438	_____	_____

The grand mean of Fisher-transformed coefficients was significantly greater than zero ($p < .01$), indicating an overall relationship between science attitude and knowledge of science for junior and senior high school students of science. No difference due to achievement level was found, but the grade difference was significant ($p < .07$). Senior high school correlations were significantly higher than junior high correlations. No achievement by grade interaction was found.

The higher correlations for senior high level students is probably due to the selection which takes place by eleventh grade. Only those students with some interest in science elect biology, physics, or chemistry, those subjects the high school sample had elected.

Conclusion

The derivation and analysis presented here allow the experimenter to perform multifactor contrasts on correlation coefficients for several populations in an experimental design. The utility of the technique is demonstrated for cross-grade level comparisons where the individual measures, such as achievement tests from which the correlations are computed, are non-comparable by themselves.

References

Anderson, G. The Assessment of the Learning Environment: A Manual for the Learning Environment Inventory. Halifax, Nova Scotia, Canada, 1971.

Box, G.E.P. Some theorems on quadratic forms applied in the study of analysis of variance problems. Annals of Mathematical Statistics, 25, 1953, 290-302.

Garibaldi, A.M. Reliability estimates for instruments used in the Minnesota Research and Evaluation Project. NSF-Minnesota Research and Evaluation Project, University of Minnesota, Minneapolis, January 1974 (mimeo) 28 pp.

Hays, W.L. Statistics. New York: Holt, Rinehart and Winston, 1963.

Lawrenz, F. Test of achievement in science, description and copies of Form I and Form II instruments. NSF-Minnesota Research and Evaluation Project, University of Minnesota, Minneapolis, March 1972 (mimeo) 23 pp.

Marascuilo, L.A. Large-sample multiple comparisons. Psychological Bulletin, 65, No. 5, 1966, 280-290.

Marks, E. ANOVA techniques for unbalanced data. Review of Educational Research, 44, No. 3, 1974, 351-364.

National Assessment of Educational Progress. Report 1, 1969-70, science: national results. Education Commission of the States, Denver, July 1970 (mimeo).

Norris, R.C. and Hjelm, H.F. Non-normality and product-moment correlation. Journal of Experimental Education, 29, 1961, 261-270.

Rao, C.R. Advanced Statistical Methods in Biometric Research. New York: John Wiley and Sons, Inc., 1952.

Searle, S.R. Linear Models. New York: John Wiley and Sons, Inc., 1971.

Thomas, C.L. and Stanley, J.C. Effectiveness of high school grades for predicting college grades of black students: a review and discussion. Journal of Educational Measurement, 6, No. 4, 1969, 203-216.

Welch, W.W. and Gullickson, A.R. A strategy for evaluating the NSF comprehensive program for teacher education. School Science and Mathematics, 73, No. 9, December 1973, 759-767.

Welch, W.W. and Pella, M.O. The development of an instrument for inventorying knowledge of the processes of science. Journal of Research in Science Teaching, 5, 1967, 64-68.

Winer, B.J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.